Customer Evaluation of Telephone Circuits with Delay

By GEORGE K. HELDER

(Manuscript received June 9, 1966)

In 1964 tests were begun in which customers making transatlantic calls, to which varying amounts of delay had been added, were interviewed after call completion to determine the circuit quality. These tests were continued in 1965 using the Early Bird satellite which some customers used extensively before being interviewed. During this period a number of different echo suppressors were also tested. Results show that the quality of telephone circuits with echo suppressors decreases with increasing delay, that previous satellite calls have no effect on the customer's opinion of his present call, and that no echo suppressor was superior for all delays although some appear to be better for the longer delays.

I. INTRODUCTION

Echo occurs on any two-wire or combination two-wire, four-wire telephone circuit. The degrading effect of this echo depends on the time it takes for the echo to return to the speaker — the delay time. For very short echo delays — on the order of a few ms — the echo is masked by the sidetone in the telephone receiver. For somewhat longer delays — on the order of a few tens of ms — the echo can be made tolerable by providing loss in the telephone circuit to reduce the level of the echo. For delays longer than this, too much loss is required and echo suppressors are used. Echo suppressors are voice operated devices placed in the four-wire portion of the circuit which insert loss in the return path to suppress the echo.

When an echo suppressor is in its suppression mode, it places a large loss in the echo path which, besides suppressing echo, prevents the speech of the second party of the conversation from reaching the first party if both are speaking at the same time, i.e., double talking. If the echo suppression is removed during double talking, the echo is also heard.

The speech mutilation and imperfect echo suppression during double talking are generally not serious for delays up to about 100 ms. For longer delays, the effects can be easily noticed by many conversants and become increasingly troublesome as the delay increases.

Many tests have been conducted during the past few years to determine the importance of these subjective effects on telephone communications over circuits with long transmission delay. The interest in this subject is related to the introduction of satellite communications systems with their inherently longer delays than terrestrial circuits. Results previously reported have been those gathered in laboratory tests.*

Late in 1963 a program was instituted to obtain the reaction of real users of circuits with different amounts of delay. Initially, the quality of circuits from the United States to Europe over cable and the *Telstar*® and Relay satellites was compared by placing regular commercial telephone calls over circuits on these media. The users of the circuits were interviewed after the completion of the calls and asked whether they had had any difficulty in conversing and asked to rate the call Excellent, Good, Fair or Poor. It was soon evident that the problems of establishing telephone circuits over the experimental satellites were producing quality differences which were confounding the desired test results.

In early 1964, at the request of the Federal Communications Commission and with the advice and knowledge of the National Aeronautics and Space Administration and the Communications Satellite Corporation, a new series of tests was begun. These tests, conducted in cooperation with the telephone administrations in the United Kingdom and France, resulted in some 3000 interviews of customers who completed calls to London and Paris over cable circuits to which varying amounts of delay had been added.

This series of tests was continued in 1965 on circuits over cable and the Early Bird satellite. 4000 more interviews resulted. The telephone administrations in Germany and Italy, in addition to those in the United Kingdom and France, participated in these tests.

This paper discusses the tests conducted in 1964 and 1965.

II. 1964 TESTS

2.1 *Test Description*

From January 27 to April 24, 1964, two non-TASI cable circuits from New York to Paris and two from White Plains to London were equipped

* See, for example, Refs. 1 and 2.

with special echo suppressors and added delay.* The special echo suppressors were used in place of the 1A and, in London, the 6A echo suppressors regularly used on the circuits. A record-reproduce device inserted delay in the outgoing path from the USA as shown in Fig. 1. Four values of total round trip circuit delay were used: 90 ms, 300 ms 600 ms, and 800 ms.† (A small number of interviews were also made using 1000 and 1200 ms delay.)

Two experimental echo suppressors, designated the B and the L, and modified British 6A echo suppressors were used on the White Plains to London circuits; the B, L, and modified Bell System 1A echo suppressors were used on the two circuits from New York to Paris. The echo sup-

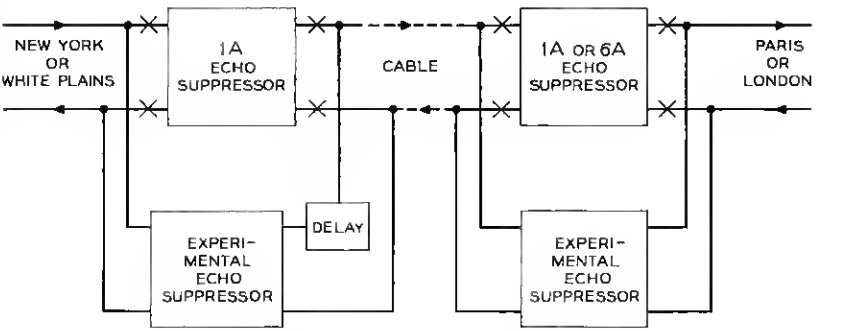

Fig. 1 — Placement of echo suppressors and delay.

pressors were rotated on the circuits so as to prevent any differences in individual circuit quality from biasing the results.

The four circuits had essentially the same transmission characteristics such as bandwidth, noise and transmission loss. The 3 dB down points of the bandwidth were at about 300 and 3100 Hz. The average noise level was about 42 dBrnC at the 0 transmission level point.‡ The circuits

* TASI, which stands for Time Assignment Speech Interpolation, is a technique used to provide more cable channels by time sharing. The circuits used in the 1964 tests were not derived in this manner.

† The 90 ms minimum value is the sum of the 70 ms inherent round trip delay of the cable plus 20 ms, which is the minimum delay introduced by the Echo-Vox. The Echo-Vox delay units were used on all calls including the control, minimum delay, calls. It can be shown that providing a lumped delay equal to the round trip delay in one direction of transmission only is, from the point of view of the speakers and the echo suppressors, equivalent to splitting the delays equally between both directions.

‡ dBrnC is dBrn as measured with a Western Electric 3A Noise Measuring Set with C Message Weighting.

were given special access codes* and were used as first choice circuits for outgoing calls from New York City to London and to Paris between the hours of 9:00 a.m. and 2:00 p.m., New York time, on weekdays. The traffic placed over the circuits was representative of the total traffic occurring during the test hours.

The operators using the circuits had no knowledge of the echo suppressors or delays being used, and they were given no special instructions other than those on the use of the specially coded circuits. The operators noted on the call ticket the access code used in addition to the normal information on party identification, call timing, etc.

2.2 *Echo Suppressor Description*

The pertinent parameter values of the echo suppressors used in the test are shown on Table I. The echo suppressor designated 1 is the modified 1A; 6 is the modified 6A.

Of the four echo suppressors used, the modified 6A and 1A echo suppressors were very similar. Both combined the suppression control and differential action into one circuit.† This does not allow a loud talker to remove suppression caused by a weak talker until the suppression hangover has expired. The modification in the 1A consisted of a 10 dB higher sensitivity and increased suppression loss. The modified 6A had a changed sensitivity frequency characteristic. This change was observed by the British Post Office (BPO) to produce suppression for lower level speech sounds with no increase in operation on noise.

The L echo suppressor, an experimental model of a United States manufacturer, was distinguished by its action during double talking. When this occurs the suppression was removed immediately, and 6-dB attenuation was inserted in both the receive and transmit paths to reduce the echo which was unsuppressed at this time. The hangover on the break-in circuit was about 30 ms. Another unique feature of this echo suppressor was the hysteresis effect present in the suppression and break-in circuits. The signal level required to release from the suppression mode was about 9 dB below that required to operate.

The B experimental echo suppressor, a Bell Laboratories model, was similar to the L in that suppression is removed immediately when double talking was detected. It was, however, unique in the long break-in hang-

* An access code is a three digit number usually associated with one country. When keyed by the overseas operator, it directs the call through the gateway switching machine to the group of circuits to that country. The special access codes here directed the calls to selected subgroups of the circuits to London and Paris.

† For a more complete description of the modified 1A and the B echo suppressors, see Ref. 3.

TABLE I — CHARACTERISTICS OF THE ECHO SUPPRESSORS IN THE 1964 TESTS (NOMINAL OR TYPICAL MEASURED VALUES)

Suppression	Echo Suppressor			
	1	6	B	L
Operate sensitivity*	−41	−26	−31	−32
Release sensitivity	−44	−29	−33	−41
Pick-up time†	10	4	8	1
Hangover time	50	50	50	30
Bandwidth	Peaked at 1000 Hz, 20 dB down at 500 and 2000 Hz	Peaked at 3500 Hz, 7 dB rise from 1000–2000 Hz	Flat over the voice band	Flat over the voice band
Break-In				
Operate sensitivity	−41	−26	−26	−36
Release sensitivity	−44	−29	−27	−41
Pick-up time	50‡	50‡	10	1
Hangover time	20	50	200	30
Bandwidth	Same as Suppression			
Loss During Double Talking				
1	None			
6	None			
B	6-dB attenuation in transmit path. Speech compressor in the receive path with 0-dB attenuation for a −40 dBm signal and 18-dB attenuation for a 0-dBm signal at the 0 transmission level point.			
L	6-dB attenuation in both the receive and transmit paths.			

* All sensitivity values are given in terms of the power level of 1000-Hz signal in dBm at the 0 transmission level point.

† All times are in msec and are measured using a 1000-Hz signal level 3 dB above the operate sensitivity.

‡ In the presence of a suppressing signal in the receive path, break-in does not occur until the suppression hangover time expires.

over time (200 ms) and in the insertion of a speech compressor during double talking to reduce the echo.

2.3 *Interviewing*

The method used to determine the subjective evaluation of circuit quality was the customer interview. The parties using the circuits were called back and interviewed by specially trained personnel in New York, London, and Paris after call completion. Fig. 2 gives the form of the questionnaire.* Every effort was made to use the same or equivalent forms and techniques in New York, London and Paris. The interviewers

* Fig. 2 is the form used in the 1965 tests. The wording of the questions on the 1964 form was essentially the same.

No. [ ][ ][ ][ ][ ] (1 2 3 4 5) Conv. Start Time [ ][ ][ ][ ] (6 7 8 9) Elapsed Time [ ][ ] (10 11)

Time First Attempt Starts [ ][ ][ ][ ] (12 13 14 15) Time Interview Starts [ ][ ][ ][ ] (16 17 18 19)

1. I am Mrs. _____ calling from the telephone company. We are making a study of our overseas connections. Do you have a moment to answer a few brief questions? Yes ☐ 20-1 No ☐ -2

2. Our records show that you made an overseas call to (the United Kingdom, Paris, Germany) a short while ago. Is that correct? Yes ☐ 21-1 No ☐ -2

3. Did you or the person you called have any difficulty talking or hearing over that connection? (Interviewer: If difficulty, probe and record verbatim - distinguish between called and calling parties.) No Difficulty ☐ 22-1 Some Difficulty ☐ -2

_____ Near [ ][ ][ ][ ][ ] (23 24 25 26 27) Far [ ][ ][ ] (28 29 30)

4. Which of these four words comes closest to describing the quality of that connection: Excellent, Good, Fair or Poor?

Excellent ☐ 31-1 EX + GD ☐ -2
Good ☐ -3 GD + FR ☐ -4
Fair ☐ -5 FR + PR ☐ -6
Poor ☐ -7 It varies ☐ -0

5. About how many times a month do you talk on the telephone with someone in (the United Kingdom, Paris, Germany)? (Interviewer: If first call, skip to Q.7)

Calls per month [ ][ ] (32 33)

Other Period Stated _____

6. On your usual call to (the United Kingdom, Paris, Germany) which of these four words comes closest to describing the quality of that connection: Excellent, Good, Fair, or Poor?

Excellent ☐ 34-1 EX + GD ☐ -2
Good ☐ -3 GD + FR ☐ -4
Fair ☐ -5 FR + PR ☐ -6
Poor ☐ -7 It varies ☐ -0

7. Do you have any other comments you would like to make about your overseas telephone service? Yes ☐ 35-1 No ☐ -2

_____ [ ][ ][ ][ ][ ][ ][ ][ ][ ][ ] (36 37 38 39 40 41 42 43 44 45)

8. Thank you very much for your help.

Accent: None ☐ 46-1 Slight ☐ -2 Heavy ☐ -3 Interpreter ☐ -4

Reasons for non-completion of interview [ ][ ][ ][ ] (47 48 49 50)

Was any person that answered irritated? Yes ☐ 51-1 No ☐ -2

Time Contact Ended [ ][ ][ ][ ] (52 53 54 55) Number of Interviewer [ ] (56) Number of Coder of Question 3 [ ] (57) No. of Coder of Question 7 [ ] (58)

Circuit Condition [ ] (59) Exposure ☐ 60-1 ☐ -2

Fig. 2 — The interview form used in the 1965 tests.

in New York were employees of a survey firm specializing in this type of work; in London and Paris they were telephone administration personnel who were given special training in interviewing. The interviewer had no knowledge of the type of circuit on which the call was made. No person was interviewed more than once. All interviews were during normal business hours and were completed the day the call was made.

2.4 *Results*

2.4.1 *General*

The primary measures of the subjective effects of transmission delay are the answers to questions 3 and 4 of the interview which concern difficulty in talking or hearing and quality rating. These questions were asked of customers in three locations: New York City, London, and Paris. The interviews in New York City have been separated into those on calls to London through the White Plains gateway and those on calls to Paris through the New York gateway. The customers used circuits with one of four different echo suppressors: 1, 6, B, and L; and had one of four values of delay: 90, 300, 600, and 800 ms. About 60 interviews were completed for each of the combinations of city, echo suppressor and delay. A total of about 3000 interviews resulted.

2.4.2 *Effect of Delay*

The interview results, pooled over echo suppressors and cities, are shown on Fig. 3, where the percentage of interviews reporting difficulty and having a fair or poor rating are plotted.* For delays of 90, 300, 600, and 800 ms each point represents about 750 interviews. The 1000 ms point represents 175 and the 1200 ms point 75 interviews. These last were taken during a two week pilot test prior to the main test and used the L and B echo suppressors only.

It can be seen that the fair or poor rating is similar to the difficulty measure. This similarity is also present when, as below, echo suppressors and cities are separated.

The rating of transmission quality (question 4) as a function of delay is shown in Fig. 4. Data from all cities and suppressors are combined. The percentage of excellent ratings decreases almost linearly with delay while both the poor and fair ratings have steeper slopes at higher delays.

* Also shown are 1965 test results discussed later.

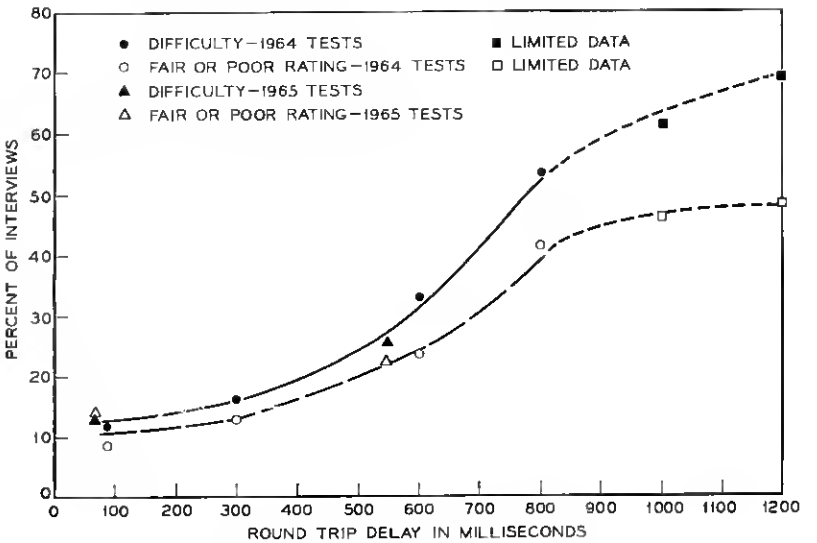

Fig. 3 — Effect of transmission delay.

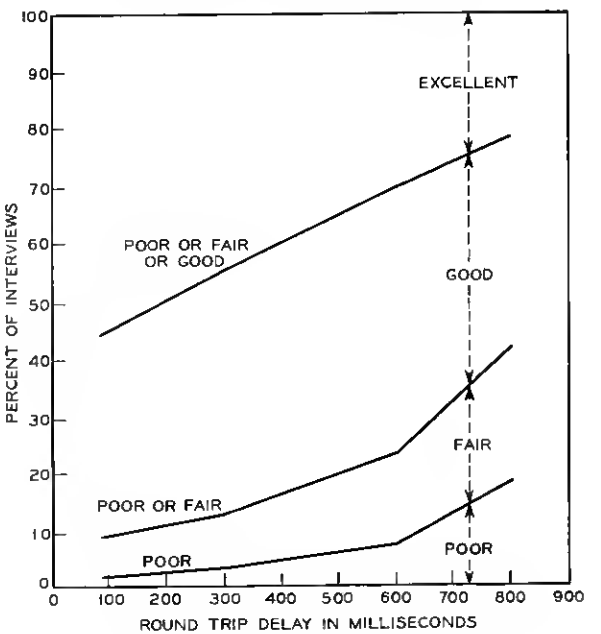

Fig. 4 — Rating of the calls.

2.4.3 *Comparison of Echo Suppressors*

Fig. 5 shows the percentage difficulty broken down for each echo suppressor separately. If the values shown in Fig. 3 are taken as the true mean, for each amount of delay, of the population of all interview results, then all but one of the values shown on Fig. 5 fall within the 95 percent confidence limits on the mean. The implication is that the differences among echo suppressors must be interpreted with caution since it is conceivable that the differences are due to chance sampling. It will be shown later that there is a more significant difference among echo suppressors in the specific types of difficulties they produce.

2.4.4 *Comparison of Cities*

The four "cities" we compare are actually three. Calls were received in London and Paris and were originated in New York City going via the White Plains gateway to London and the New York gateway to Paris. Hence, our four cities, New York (N), Paris (P), White Plains (W) and London (L).

The difficulty percentages combined over echo suppressors are shown on Fig. 6. An analysis of variance shows that, taking all delays into consideration, there is no significant difference among the cities. However, considering 90 ms only, there is a large difference between the

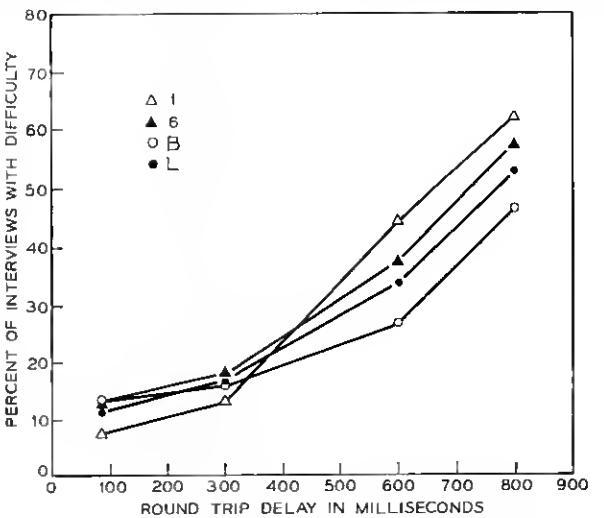

Fig. 5 — Comparison of echo suppressors.

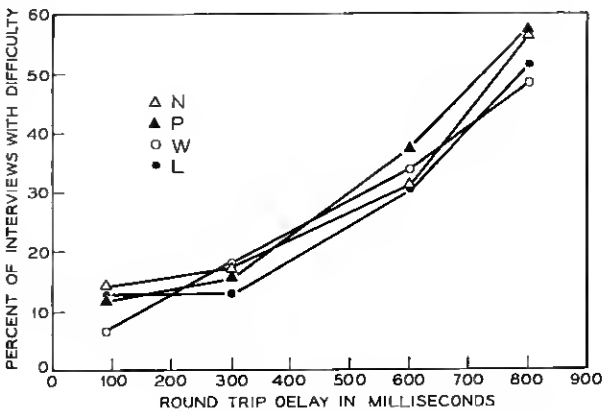

Fig. 6 — Comparison of cities.

results for New York City interviews on calls to London — W interviews, and on calls to Paris — N interviews. This difference also shows up in the usual call rating (Section 2.4.6) and in the 1965 tests (Section 3.4.3).

2.4.5 *Comments*

Those persons who had difficulty generally also described the difficulty they experienced. The difficulties, as stated in the users own words, are not always easy to classify. A three-man committee composed of persons from the Bell Telephone Laboratories, National Aeronautics and Space Administration, and Communications Satellite Corporation, without any knowledge of the echo suppressor or delay relating to each comment, classified the comments into 18 different types. The percentage, by type, of the total comments for all delays, echo suppressors and cities is shown on Fig. 7. Occasionally more than one type comment resulted from one interview.

The comments made by the users are also shown divided into comments they made about their difficulties (near end) and comments made about difficulties experienced by the other party (far end).

The relationship of comments to echo suppressors is shown on Fig. 8. The types of comments shown are those which are grouped on Fig. 7. For each comment four points are shown — one for each delay. The points represent the percentage of *interviews* which both had a "yes" answer to question 3 *and* which had a comment of the type shown, except

for the favorable comment points. In all but one case of favorable comments, no difficulty was reported and the comments were freely given.

Some things are readily apparent from Fig. 8. First, favorable comments generally decrease with delay; second, cutting comments generally increase with delay; third, echo and delay comments occur too rarely to permit meaningful interpretation. There is no clear pattern to most of the rest of the comment types. At the longer delays the B echo suppressor produces less cutting comments than the other echo suppressors. It is believed that this is due to the longer break-in hangover time and the insertion of the speech compressor during double talking.

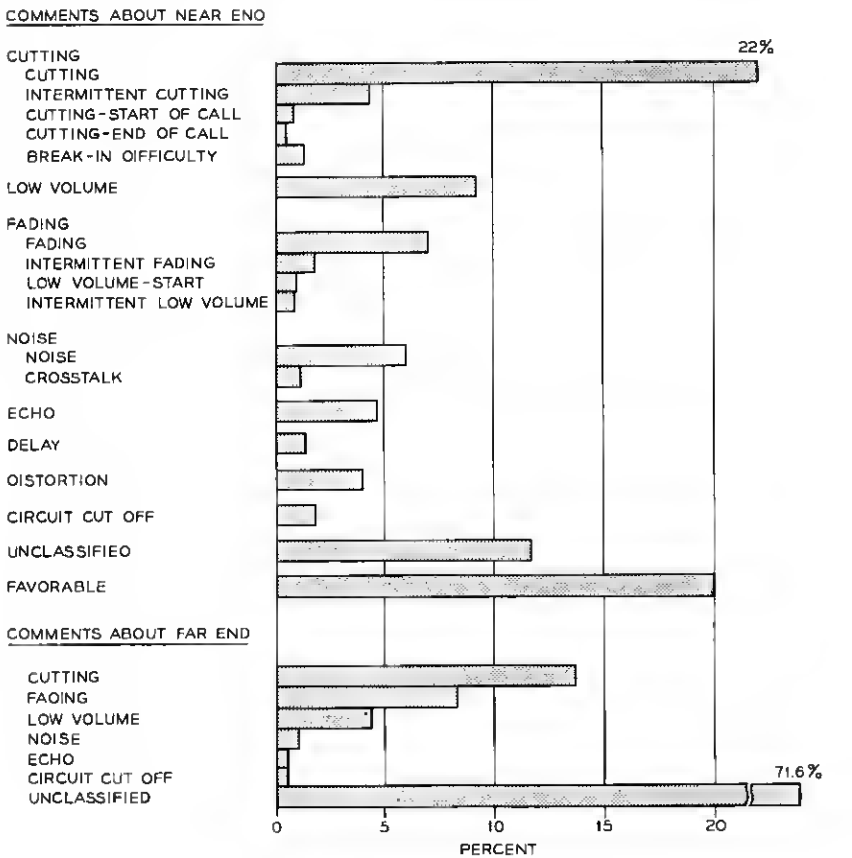

Fig. 7 — Percent of total comments by type.

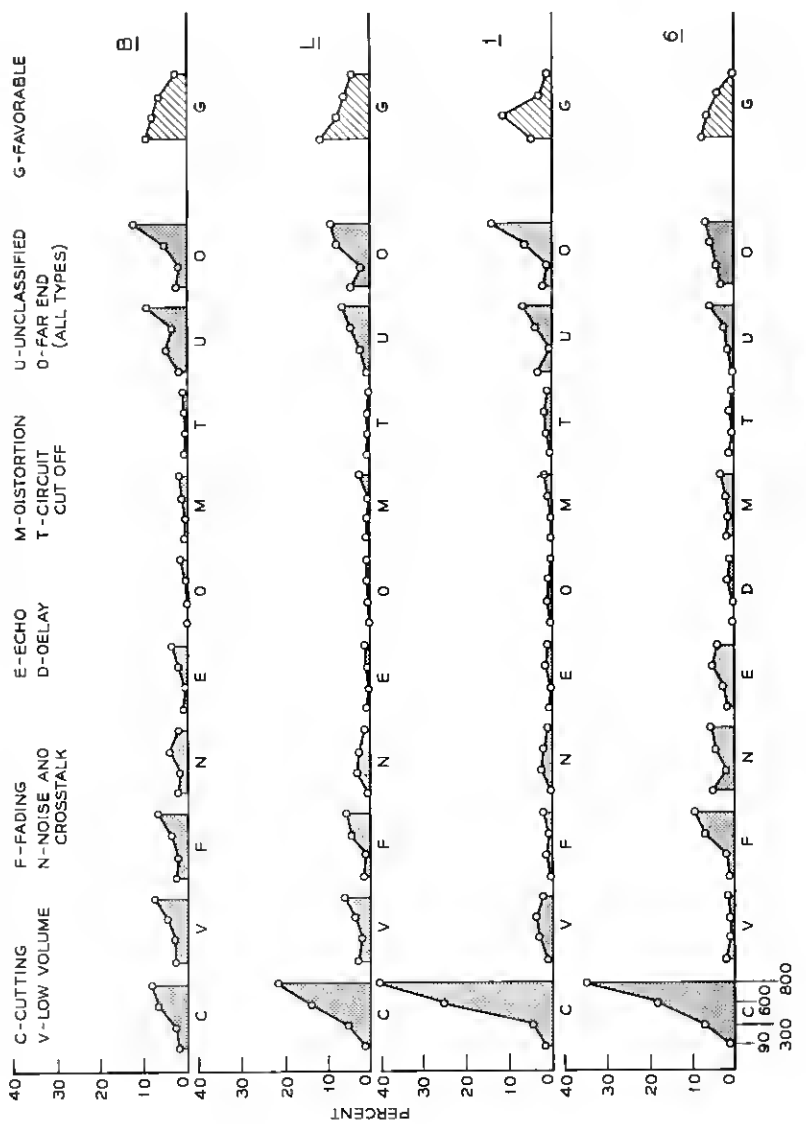

Fig. 8 — Comment types by echo suppressor.

2.4.6 *Rating of Usual Call*

Those interviewed who were not first time users of circuits to London and Paris replied to question 6 on the rating of their usual call. The results are shown on Fig. 9 pooled over cities and echo suppressors where each point represents about 700 interviews. (Also shown are 1965 results discussed later.) The usual call ratings change little with delay while the ratings of the present call, Fig. 4, are strongly affected by delay.

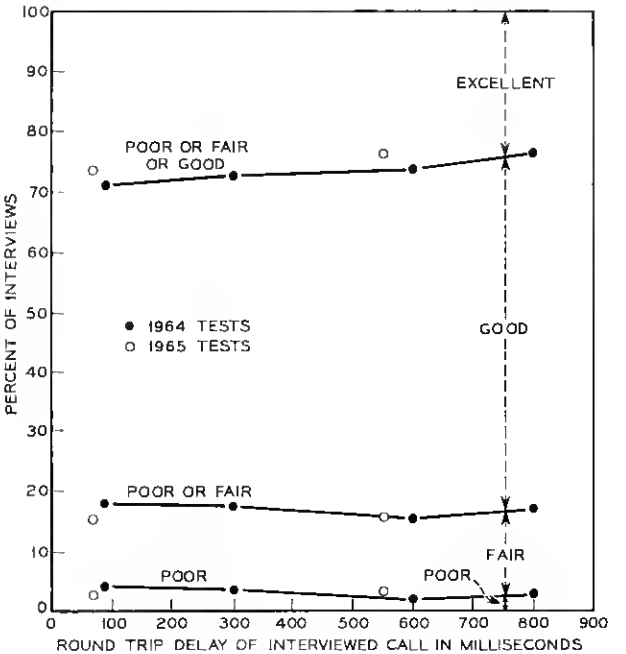

Fig. 9 — Rating of the usual call.

The average rating of the usual call is essentially independent of the average rating of the present call. Also of interest is the rating of the usual call compared with the rating given the present 90 ms call which is very similar in quality to the usual call.* The usual call ratings show fewer excellent and considerably more fair ratings than the present call. The predominant rating of the usual call is good. This may be due to an averaging by the customer of all his previous calls which can range from excellent to poor.

* The usual call is either non-TASI or TASI cable using 1A or equivalent echo suppressors.

The usual call ratings for each city separately, combined over delay and echo suppressors, are shown on Fig. 10. The differences in quality of the usual call to London and Paris for the New York City party (W and N) are consistent with the 90 ms results shown on Fig. 6. The differences between the London and Paris usual call ratings could easily be due to the differences of the population of users in the two cities.

2.4.7 *Length of Conversation*

The length of conversation was not related to the transmission delay as shown on Fig. 11. Prior to the test it was speculated that conversation time might decrease because of decreased satisfaction with the circuit quality, or might increase because of the lost information during times of echo suppressor mutilation and confused situations. Additional analysis shows no significant correlation bewteen the length of conversation and the percentage of customers reporting difficulty.

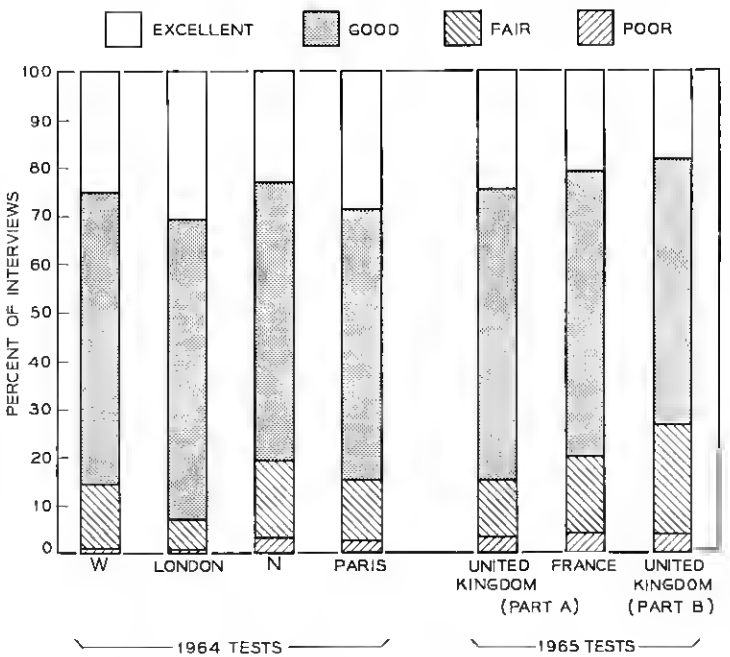

Fig. 10 — Rating of the usual call for each city.

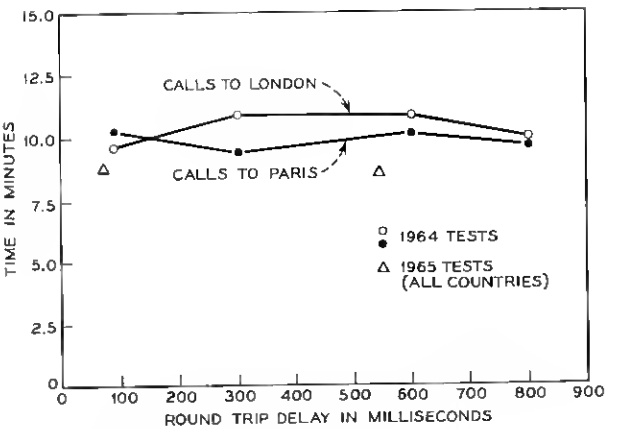

Fig. 11 — Average length of conversation.

2.4.8 *Speech Volume*

Speech volume measurements were made on most calls by craftsmen using VU meters. Some 4500 measurements indicate that the volumes for each delay were very similar. The greatest difference of averages between any two delays was 0.8 dB.

The percentage of customers reporting difficulty as a function of their volume is shown on Fig. 12 where a trend towards more difficulty for higher volumes exists (or conversely, higher volume for more difficulty).

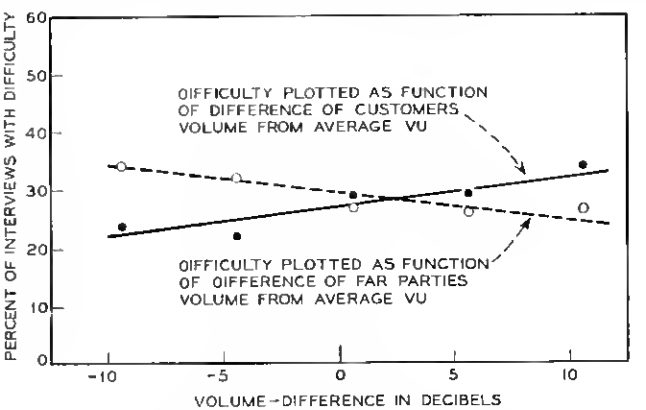

Fig. 12 — Effect of speech volume on difficulty.

Each point shown represents between 64 and 608 volume measurements. Also shown on Fig. 12 is the difficulty as a function of the received volume of the other party (not the interviewee). Here the trend is reversed, the interviewed party having less difficulty as the far party's speech volume increases. Each point represents between 182 and 480 volume measurements.

These trends might reasonably be expected. The louder one party talks, the more likely he will control the echo suppressors and cut off the interrupting speech of the far party. This also is more likely to happen if the far party is a low level talker. Additionally, this low received level itself may cause difficulty for the interviewed customer in the absence of any double talking.

2.4.9 *Interviews at Both Ends*

On 1000 calls it was possible to interview both parties to the conversation. Table II presents the percentage of calls on which the various combinations of difficulty were present.

What one can say is that if either party to the call reported difficulty there is a 50–50 chance the other party also will report difficulty, and if either party reported no difficulty, there is about a four to one chance the other will also report no difficulty. The results for Paris and London separately are very similar.

III. 1965 TESTS

3.1 *General*

In the 1964 tests the customers had only one exposure to long delay before interview. It had long been thought that frequent users of long delay circuits may react differently than those who seldom use them. It was unknown whether these heavy users would accommodate to the delay and echo suppressors or become more critical or neither. The advent of the Early Bird satellite provided the opportunity for investigation.

TABLE II — PERCENTAGE OF CALLS WITH DIFFICULTY

USA end	Europe End	
	Yes difficulty	No difficulty
Yes difficulty	15%	14%
No difficulty	14%	57%

The principal tests in 1965 (termed Part A) were aimed at determining the subjective quality of satellite circuits when:

(*i*) both satellite and cable circuits were in general use,
(*ii*) the presence of both types was common knowledge,
(*iii*) the novelty, if any, of satellite circuits had worn off, and
(*iv*) heavy users of transatlantic circuits had used the satellite extensively.

To provide a baseline for comparison, the quality of non-TASI cable calls was also evaluated and, in a later part of the test, an evaluation of cable circuits derived by TASI was also made.

Another series of tests, termed Part B, was conducted to evaluate a number of different connections, which may be typical in the future, involving satellite circuits with different types of echo suppressors at the two ends or with other echo suppressor-equipped circuits in tandem.

3.2 *Part A Test Description*

Commercial transatlantic telephone communications via the Early Bird satellite began June 28, 1965. From that date until November 12, 1965, all weekday satellite traffic to and from the United Kingdom, France, Germany, and Italy was restricted to that originating or terminating in New York City (except for that in the Part B tests, to be discussed later). A punched card record of each call was kept by each administration from whose national network the call originated. The punched card data included the names and telephone numbers of calling and called parties, test day, call timing, country identification, and customer recall of the originating operator. Copies of all punched cards were regularly forwarded to Bell Telephone Laboratories where they were combined to give the complete satellite calling history of each New York City overseas customer.

The calling history was complete for all weekday calls to and from these four countries for each New York City telephone number. Where different telephone numbers were used by one customer it was sometimes, but not always, possible to relate the calls on each number to the same customer. Calls made on weekends or holidays or to other countries having satellite circuits were not included in the call record cards.* About 50,000 call record cards were kept.

Interviewing of New York City customers began on September 7 for calls on the satellite and on non-TASI cable, and on October 29 for calls

* During the tests, 61 Early Bird circuits to Europe were in use. Two of these, to the United Kingdom, were used in the Part B tests. Of the remaining 59, 46 or 78 per cent were to the four countries involved.

ou TASI cable circuits. The interviews continued through November 19. Satellite and non-TASI cable interviews were conducted on calls from the USA to the United Kingdom, France, and Germany; TASI interviews were made on calls to the United Kingdom and France only.

3.3 *Part A Circuit Description*

Essentially the whole group of Early Bird circuits to each country was used for interview calls. The circuits were given special access codes and used for New York City traffic only.

A representative sample of cable circuits — 8 cable circuits to France, 7 to the United Kingdom, and 2 to Germany — were removed from general message use, assigned special access codes and used exclusively for New York City traffic. The cable circuits to France and the United Kingdom were in turn divided into two groups with separate access codes, and the groups were alternately made non-TASI or TASI following a weekly schedule.

The operators handled traffic on all cable and satellite circuits in the same manner and noted on the call ticket the access code used. Subsequent to the interview, the access code was used in the analysis to separate the calls into non-TASI, TASI, and Early Bird calls.

The non-TASI circuits to the United Kingdom were in the Cantat and TAT-3 cables. When derived from TASI, the circuits had equal probabilities of using Cantat, TAT-1 and TAT-3 cable channels. The non-TASI circuits to France used TAT-2, TAT-3, and TAT-4 channels which were also used for the TASI derived circuits. The two non-TASI circuits to Germany used channels in TAT-4.

Noise and loss measurements were made regularly on all channels. Those made at the New York and White Plains testboards are shown in Table III. The noise values are corrected, i.e., those which would have been present had the average loss been zero dB.

The delay at 2000 Hz on two London Early Bird circuits was measured from White Plains when operating with the ground station at Raisting,

TABLE III — NOISE AND LOSS MEASUREMENTS

	Loss		Noise	
	Average (dB)	Std. dev. (dB)	Average (dBrnCO)	Std. dev. (dB)
Early Bird	−0.26	0.98	42.8	2.8
Non-TASI circuits in test	0.17	1.16	43.0	2.0
All channels used in TASI	0.21	1.4	43.2	2.3

Germany and found to be 544 ms round trip. The round trip delay on a TAT-1 circuit was 73.5 ms and for a TAT-3 circuit 86.5 ms. The bandwidth of the Early Bird circuits at the 3 dB down point was about 170 to 3400 Hz; for the TAT-3 cable circuit it was 230 to 3200 Hz.

The normal echo suppressors were used on the cable circuits. These are the Western Electric 1A in the USA and France and Germany, and the GPO 6A in the United Kingdom.* The Western Electric 2A echo suppressor was used at both ends of all Early Bird circuits. Its operation is similar to the B echo suppressor without the 6-dB attenuation in the transmit path.

3.4 *Part A Results*

3.4.1 *Effect of Exposure to Satellite Delays*

The pooled results of interviews made in New York City on calls to the United Kingdom, France, and Germany are shown on Fig. 13. A heavy user is here defined as one who made or received six or more satellite calls prior to being interviewed, including the call leading to the interview if the interview was made after a satellite call. Statistical tests indicate that the differences between the interview results for light and heavy users after both cable and satellite calls are not significant at the 0.05 level. This applies to both the percent difficulty and percent fair or poor.†

Table IV gives a more detailed breakdown of satellite exposure prior to interview. The division between light and heavy users at six satellite calls is arbitrary. On the basis of other divisions, however, one would also probably conclude that there is no significant change in the results with increased exposure.

In all further presentations of results, heavy and light user interviews will be pooled.

3.4.2 *Comparison of Cable And Satellite Circuits*

The interview results for non-TASI, TASI and Early Bird circuits pooled over all countries and over heavy and light users are shown on Fig. 14. The differences between non-TASI and TASI cable are not statistically significant; those between the cable and Early Bird are highly significant.

* See Ref. 3 for a description of the 1A echo suppressor. The 6A is very similar.

† The largest difference is in fair or poor scores on satellite circuits. This difference would have to be 6.4 percent to be significant at the 0.05 level.

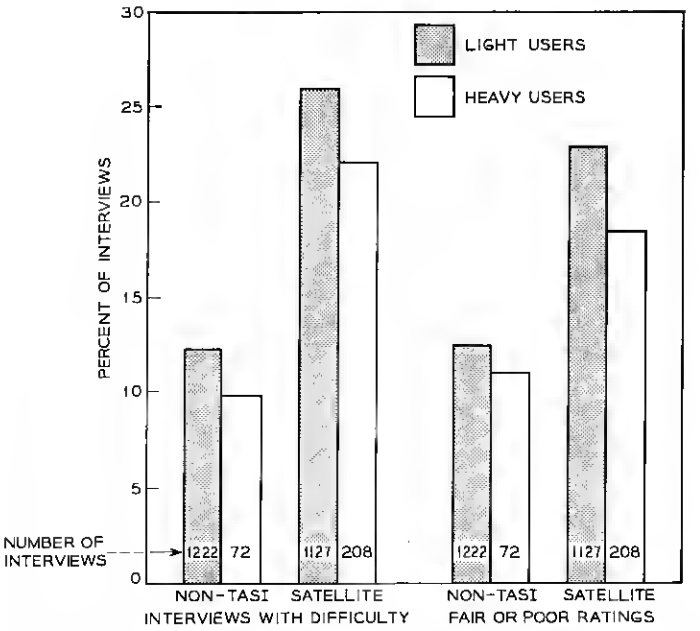

Fig. 13 — Comparison of light and heavy users of satellite circuits.

The rating breakdown is shown on Fig. 15. In addition to the breakdown for all calls, the ratings for those calls on which difficulty was or was not reported is also shown. This was done to see if there was perhaps a difference in the ratings given to calls with difficulty whether on cable or satellite. It was thought that there could be a difference in the severity of the difficulties experienced on the two types of circuits which would

TABLE IV — INTERVIEW RESULTS

No. of satellite calls prior to interview	Cable Interviews			Satellite Interviews		
	No. of interviews	Percent difficulty	Percent fair or poor	No. of interviews	Percent difficulty	Percent fair or poor
0–1	987	12.1	12.7	712	25.3	21.5
2–3	181	13.3	11.6	330	26.4	25.0
4–5	54	13.0	14.8	85	27.1	25.9
6–7	25	12.0	8.0	60	21.7	20.0
8–9	19	10.5	15.8	58	19.0	16.4
10–11	11	9.1	18.2	30	23.3	20.0
12–13	7	14.3	14.3	20	25.0	15.0
14–15	4	0	0	7	0	14.3
16 & Over	6	0	0	33	30.3	21.2

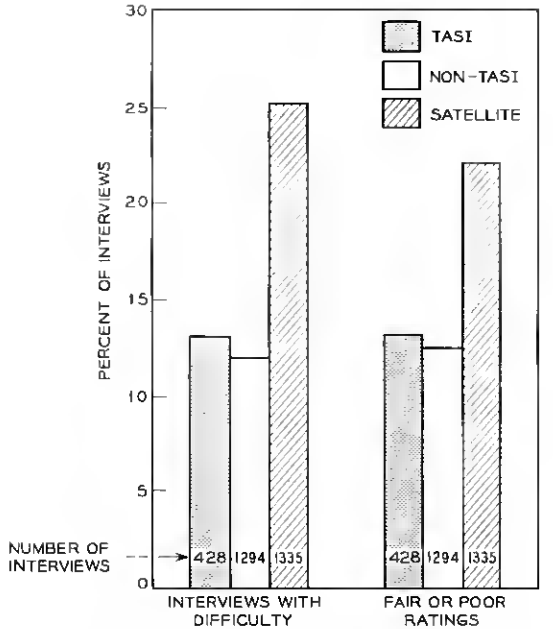

Fig. 14 — Comparison of cable and satellite circuits.

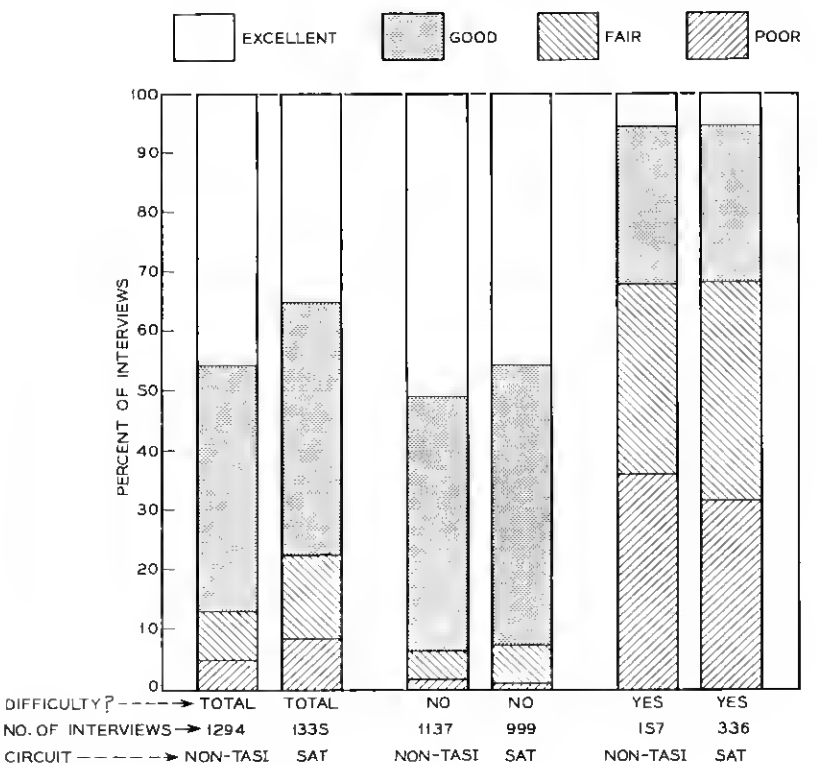

Fig. 15 — Rating of the calls.

show up in the ratings. The similarity of the ratings of the two types of circuits within both 'yes' and 'no' categories of difficulty suggests that there is no difference.

3.4.3 *Comparison of Countries*

The interview results broken down by country of destination are shown on Fig. 16. The following differences between the United Kingdom and France are statistically significant:

(*i*) TASI fair or poor ratings (significant at the 0.01 level).
(*ii*) Non-TASI difficulty (significant at the 0.01 level).
(*iii*) Non-TASI fair or poor (significant at the 0.05 level).

Between the United Kingdom and Germany only the non-TASI diffi-

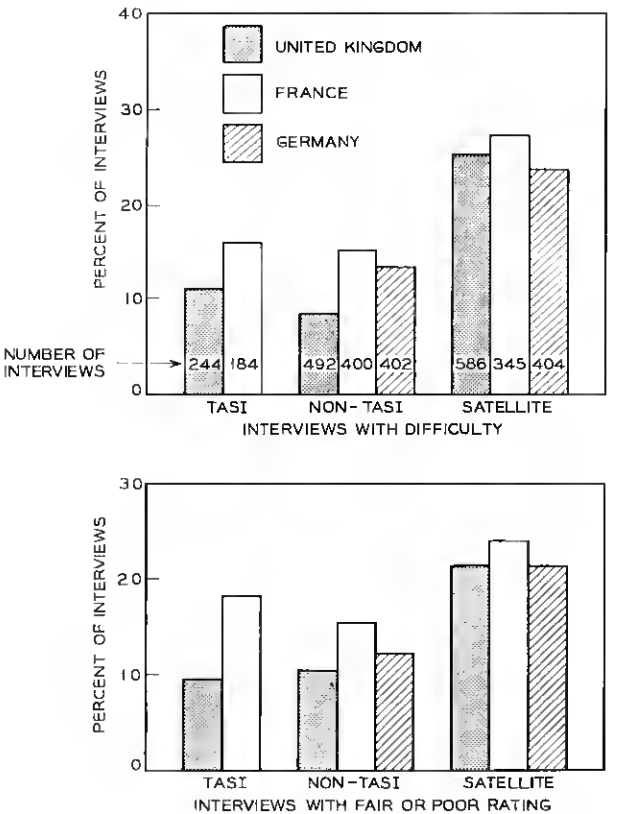

Fig. 16 — Comparison of countries.

culty difference is significant at the 0.05 level. No other differences between countries shown on Fig. 16 are significant. In Section 2.4.4 the differences between interviews on calls to London and Paris for low delays are mentioned.* The reason for this consistent difference is not known.

3.4.4 *Comparison of Ground Stations*

During the test period, the ground stations at Goonhilly Downs, Pleumer-Bodou and Raisting worked with Andover on a weekly rotating basis. The ground stations fed the international gateways in a triangular interconnecting circuit arrangement shown simply on Fig. 17. The ground

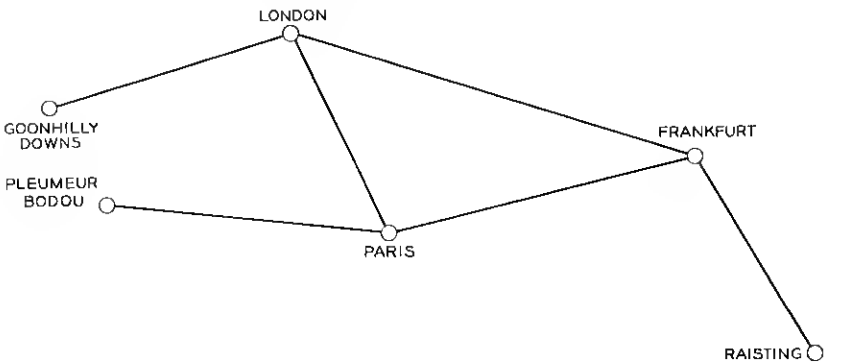

Fig. 17 — Europe Early Bird network.

station at Goonhilly alternately distributed the continental traffic through Frankfurt and Paris via London.

The interview breakdown by ground station, pooled over countries, is shown on Fig. 18. None of the differences are significant. A further analysis for each country separately also fails to yield significant differences, e.g., the United Kingdom results are the same for all ground stations.

3.4.5 *Customer Comments*

Prior to the 1965 tests, the comments made in the 1964 tests were analyzed to determine the most descriptive words. These descriptive

* In the 1965 tests 84 percent of the interviews on United Kingdom circuits were on calls to London and 90 percent of the interviews on France circuits were on calls to Paris.

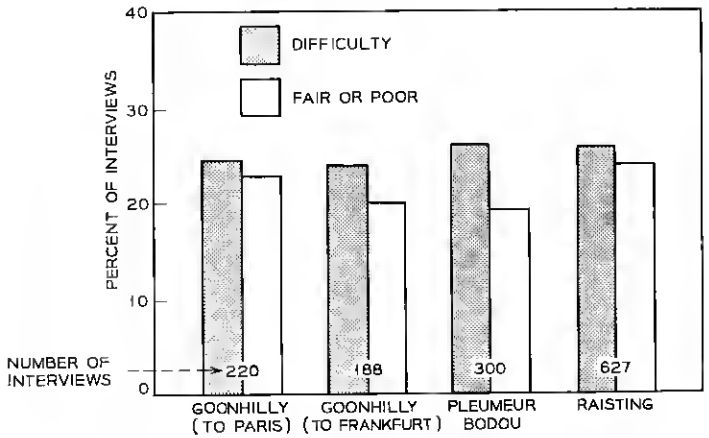

Fig. 18 — Comparison of ground stations.

words were then related to nine categories of difficulty and one favorable category, and a dictionary was prepared. The categories were a collapsed version of those used for the 1964 test as shown below.

Dictionary coding	1964 cable test coding
1 Cutting	Cutting on and off
	Cutting at start
	Cutting intermittent
	Break-in difficulty
	Cutting at end
2 Low Volume	Low volume
	Low volume at start
	Low volume intermittent
3 Fading	Fading
	Fading intermittent
4 Noise	Noise
5 Echo	Echo
6 Distortion	Distortion
7 Delay	Delay
8 Crosstalk	Crosstalk
9 Unclassified	Unclassified
	Circuit cut off
10 Favorable	Favorable

The interview supervisor, a senior interviewer, coded the comments made in the 1965 test into these ten categories using the dictionary as a guide. Fig. 19 presents the frequency of each type of comment in terms of the percent of all interviews having comments of the type shown. These include comments relating to the difficulties encountered by the near (interviewed) party and by the far party as described by the near party. In some cases interviews had comments classifiable into more than one category and occasionally both the near and far parties had

the same type of difficulty. About five percent of the customers who stated they had no difficulty, and were recorded as such, also made a classifiable comment about difficulty on the circuit. These data are included. On about two percent of the interviews, a favorable comment was made about transmission for one party along with difficulty comments about transmission for the other party.

It can be seen that in general the differences between the percentages shown are small and not consistent when comparing non-TASI and TASI. When comparing cable and satellite circuits the differences are consistent and, in many cases, statistically significant. The troubles experienced by the users of the satellite circuits appear to be caused in large part by the action of the echo suppressor as evidenced by the large increase in the cutting, fading, and echo comments. The delay *per se* of satellite circuits does not appear to be noted by the customers, but rather it is the delay in combination with the necessary echo suppressors.

3.4.6 *Rating of the Usual Call*

The customer's rating of the usual call was asked in each interview in order to be able to interpret his rating of the present call relative to

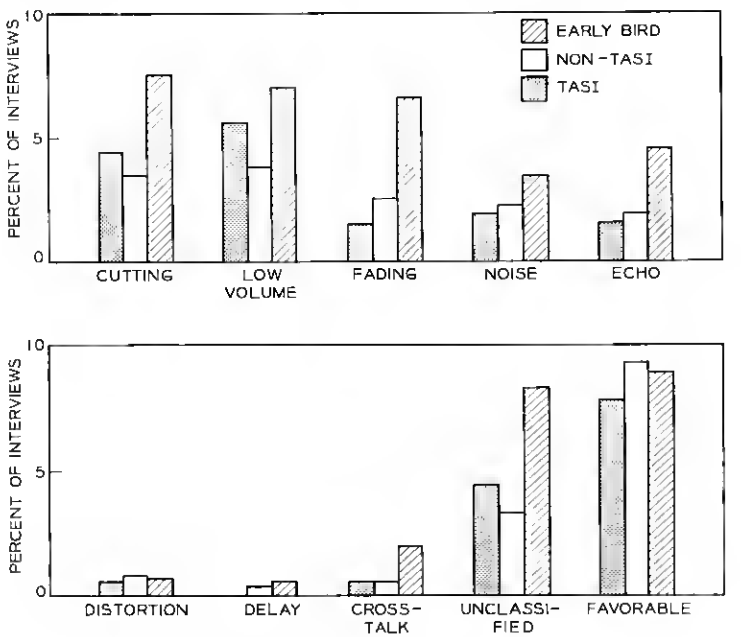

Fig. 19 — Customer comments.

his rating of his usual call to the same city. The usual call ratings for interviews on cable and satellite are shown on Fig. 9 where it is seen that they are very similar to those of the 1964 tests.

The relation between present and usual call ratings is shown on Fig. 20 in terms of the percentage of people who rated the present call one, two or three rating steps better or worse than their usual call. The satellite interviews have more present calls rated one, two or three steps poorer than the usual call than do the cable interviews. On both satellite and cable interviews, however, there are more cases in which the present call is rated better than the usual call than cases where the present call is rated worse than the usual call.

The usual call ratings for each country separately, combined over cable and satellite interviews, are shown on Fig. 10. A comparison of the W and United Kingdom and the N and France bars shows good agreement between the 1964 and the 1965 Part A tests and suggests that the population of customers and the usual call quality are the same in both tests.

As mentioned in Section 2.4.6, the respondents tend to give their usual (presumably cable) call a predominant rating of good and this

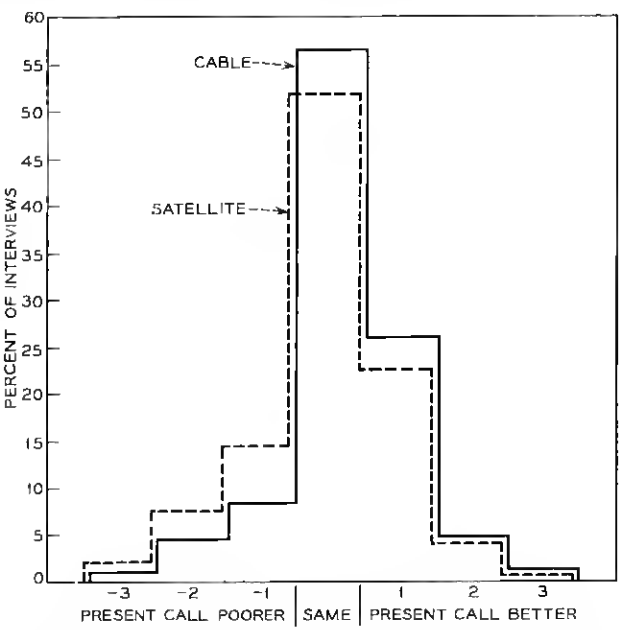

Fig. 20 — Rating of the present call expressed as the difference in rating steps from the usual call.

differs from the rating given present cable calls from essentially the same population of calls. This difference in ratings has been noted in several other tests. Because of this effect it is not appropriate to compare ratings of the usual calls and ratings of present calls using different facilities to determine relative transmission quality without compensating for the above described tendency.

3.4.7 *Length of Conversations*

As shown on Fig. 11, the average conversation length was essentially the same for cable and satellite calls.

3.4.8 *Time of Conversation*

Almost all interviews were made on calls which occurred between the hours of 8 a.m. and 2 p.m. New York time. These hours include the busiest traffic hours and an hour or so of lighter traffic before and after the busy hours. The assumption that the rate of customer difficulty increases with traffic density was tested. The following positive significant correlations with traffic density were found:

(*i*) Difficulty and fair or poor on United Kingdom TASI circuits.
(*ii*) Fair or poor on United Kingdom non-TASI circuits.
(*iii*) Fair or poor on France non-TASI circuits.
(*iv*) Difficulty on France Early Bird circuits.
(*v*) Fair or poor on German non-TASI circuits.

The reason for these correlations is not known. The correlations are not consistently associated with any country or type of circuit.

3.5 *Part B Test Description*

The Part B tests were conducted on two Early Bird circuits to London equipped with special echo suppressors. Traffic between the United Kingdom, excluding London, and 21 numbering plan areas surrounding but excluding New York City was placed on these circuits and the parties were interviewed after call completion. Confinement to the 21 NPA's prevented the inclusion of tandem echo suppressors in the United States Direct Distance Dialing (DDD) network.

3.6 *Part B Echo Suppressors*

The echo suppressor combinations used on the circuits are shown on Fig. 21. The 2A and 6A echo suppressors have already been described. (The 6AF is a full version of the 6AS used on cable circuits.) When used as split (6AS) echo suppressors, condition 6, 2-ohms delay was inserted between the two 6AS echo suppressors and −50 dBmO (psophometric) of noise was injected.

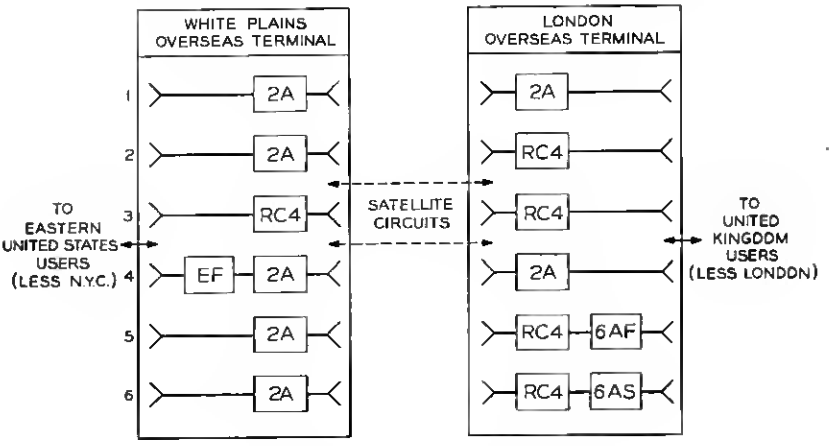

Fig. 21 — Part B circuit combinations.

The RC4 is a British Post Office experimental echo suppressor designed for long delay circuits. Its distinguishing features are a shaped frequency response in the control circuit path (like the modified 6A of Table I), a break-in hangover of 300 ms and 6 dB of loss inserted in the receive path during double talking.

The EF echo suppressor, a Bell Laboratories experimental model, is a full echo suppressor proposed as a replacement for the present Western Electric 1A in most domestic applications. It is much like two 2A echo suppressors back to back without speech compressors in the transmission paths.

3.7 *Part B Test Results*

3.7.1 *Comparison of Echo Suppressor Conditions*

The Part B results of interviews in the United States only are shown on Fig. 22. An analysis of variance shows that, taken as a group, there are no significant differences among the six conditions.

The average percentage difficulty for all conditions in the Part B tests was 42.4 and the average percentage fair or poor was 27.9. These values are close to those for condition 1 which used 2A echo suppressors only. It will be noted that these percentages are considerably different from those in the Part A tests for calls to the United Kingdom which were 25.1 percent difficulty and 21.6 percent fair or poor. The reason for this is not known. The populations of users were different in that the

Part A customers were in New York City and conversed, in most cases, with someone in London, while the Part B customers were in the eastern United States and conversed with someone in the United Kingdom excluding London. This may result in somewhat poorer circuit quality. Speech volume readings were made at the London gateway by the British Post Office, and the average transmitted volume to New York was 2.3 dB lower in the Part B tests than in Part A. As discussed in Section 3.4.8, this contributes to poorer quality. The ratings of the usual calls are also lower in Part B than in Part A.

3.7.2 *Rating of the Usual Call*

The usual call ratings are shown on Fig. 10 along with those Part A calls to the United Kingdom. This shows the higher percentage of fair ratings and lower percentage of excellent and good ratings in the Part B tests. No interviews were made on cable calls in the Part B tests so no comparison can be made between usual and present calls.

3.7.3 *Customer Comments*

One way of comparing the types of difficulties produced by the combinations is as follows. Combinations 1, 2, and 3 involve no tandem echo suppressors and the comments are therefore pooled. Combination 4 has

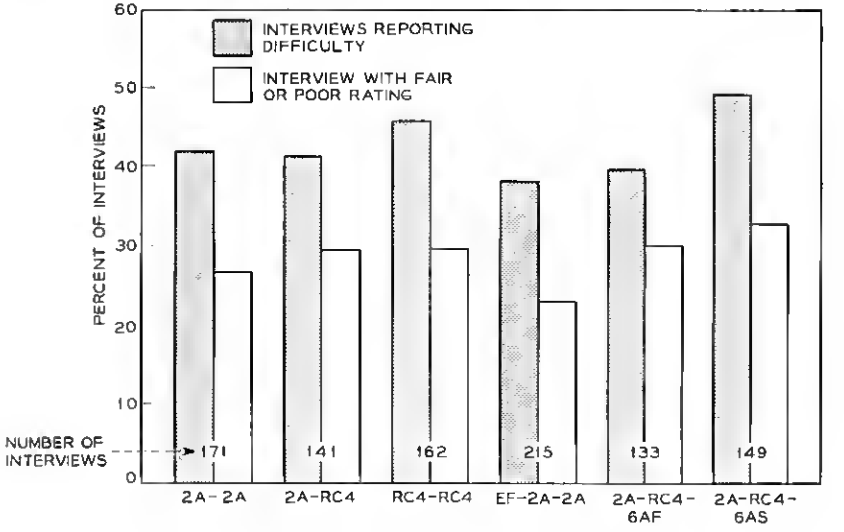

Fig. 22 — Comparison of circuits.

a tandem suppressor at the near end. Combinations 5 and 6 have tandem suppressors at the far end, and these comments are pooled. Taking the sum of all customer comments (excluding favorable) for each of the three classes of circuits above, one can then determine the percentage of the difficulty comments falling into each category, rather than the percentage of interviews having comments in each category. This has been done and is shown on Fig. 23. As would be expected, the tandem echo suppressor combinations generally produce a greater proportion of cutting and fading comments and fewer echo and crosstalk comments. (A customer comment about fading can relate to echo suppressor action, and echo is sometimes described by the customer as "other voices on the line," i.e., crosstalk.)

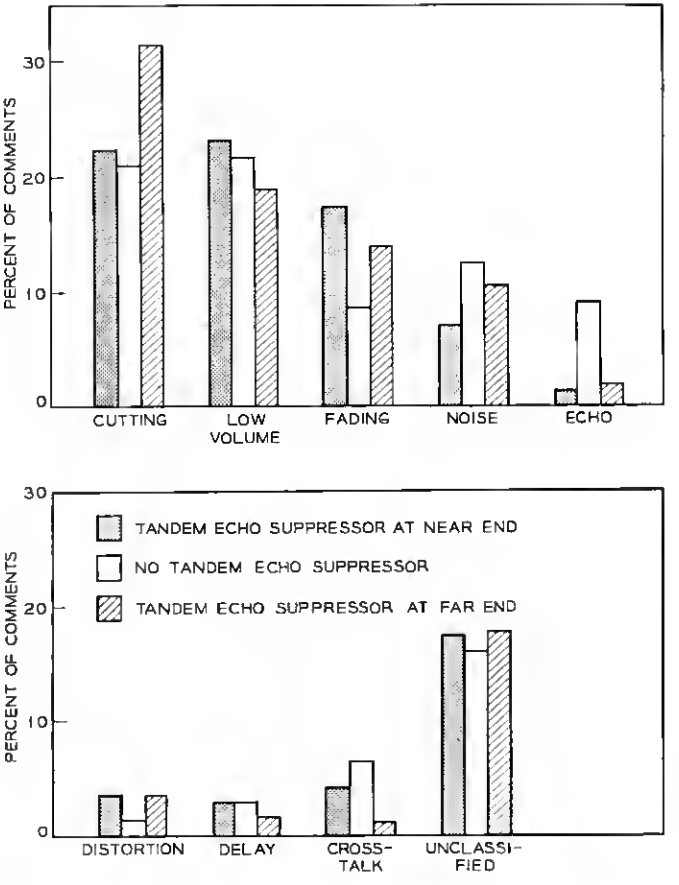

Fig. 23 — Comment breakdown.

IV. INTERVIEWING

4.1 *Technique*

Most of the interviewers in 1964 and 1965 were housewives who interviewed on a part-time basis. They had previous training in interviewing by their employer and were given two days of specialized training for these tests.

The interview took an average of 2.1 minutes to complete. About 0.6 percent of the customers in 1965 expressed some irritation about being interviewed. Although no count was taken, a sizable percentage of interviewees expressed some pleasure about being interviewed.

In 1964, the average time from call completion to interview completion was considerably less than one hour and in many cases only a few minutes. During the much more complex 1965 tests, the average time was greater than one hour. Some 81 percent of all interview attempts in 1964 resulted in a completed interview and this high percentage is due mainly to those interviews conducted within a few minutes. In 1965, this rate dropped to 65 percent.

The main reasons for non-completion of interviews appear in Table V.

4.2 *Comparison of Interviewers*

Five interviewers conducted most of the interviews in the 1965 Part A tests (one interviewer made $\frac{2}{3}$ of the interviews in the Part B tests and no Part A interviews). The Part A interviewers are compared on Fig. 24. Generally, the differences among interviewers are small compared with the differences among circuits, although in one case the cable ratings were lower for the cable than the satellite — the fair or poor ratings for interviewer number 4.

V. COMPARISON OF 1964 AND 1965 TEST RESULTS

The differences between the 1964 and 1965 tests, particularly in echo suppressors, make comparisons difficult. However, as discussed below,

TABLE V — REASONS FOR NOT COMPLETING INTERVIEWS

No answer	23%
Customer busy	66%
Customer refused all interviews	1%
Language barrier	1%

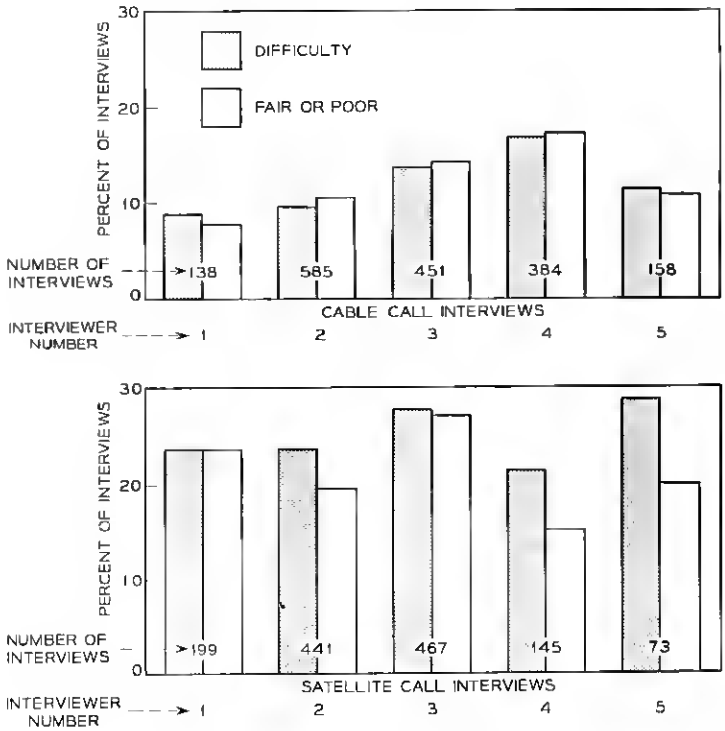

Fig. 24 — Comparison of interviewers.

those comparisons which seem most reasonable show the results of the two tests to be very similar.

In 1965, the Part A tests used the 2A echo suppressor, which is very similar to the B echo suppressor of the 1964 tests, and the interviews were in New York City only. The 1964 interviews in New York City only using the B echo suppressor are shown on Bar 1 on Fig. 25 while the 1965 results are shown on Bar 5. (The 1964 test results have been linearly extrapolated to the 545 ms delay of Early Bird. There is no extrapolation of the low cable delays.) Comparison of Bars 1 and 5 for both cable and satellite delays shows elose similarity.

Since some of the 1965 interviews are on calls to Germany, these have been removed in Bar 4 to provide interviews to the United Kingdom and France only. The same comparisons can be made.

In the 1964 tests there was no consistent difference among the echo suppressors, and to increase the sample size the New York City inter-

views for all suppressors have been combined in Bar 2. There were also no significant differences among cities. Thus, Bar 3 represents all interviews in the 1964 tests (at the delays shown).

The data represented by each bar is summarized below:

Bar number 1 represents 1964 interviews in New York only on calls to London and Paris using the B echo suppressor.

Bar number 2 represents 1964 interviews in New York only to London and Paris using all echo suppressors.

Bar number 3 represents 1964 interviews in all cities using all echo suppressors.

Bar number 4 represents 1965 New York interviews on calls to the United Kingdom and France only.

Bar number 5 represents 1965 New York interviews on calls to the United Kingdom, France and Germany.

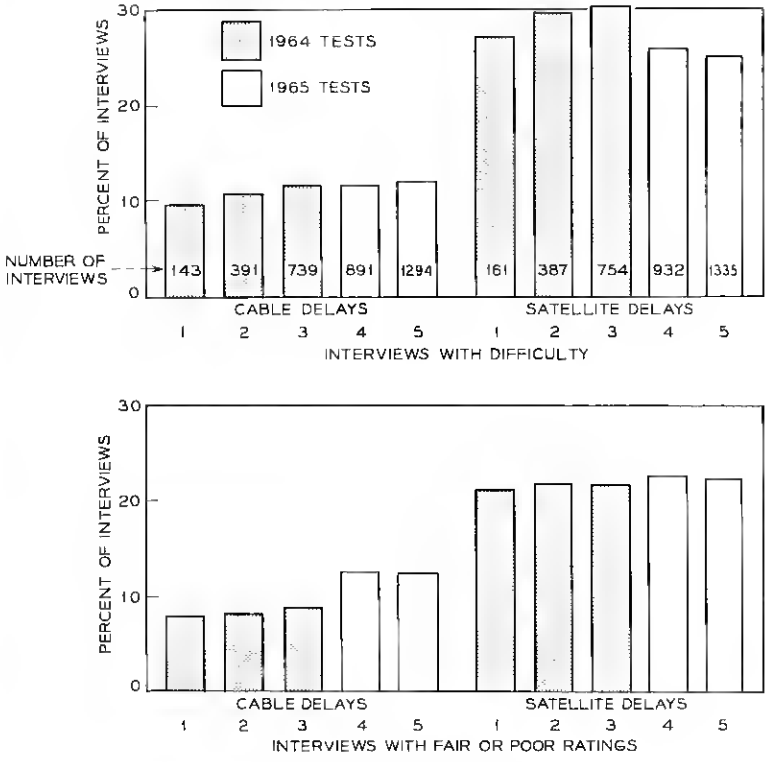

Fig. 25 — Comparison of 1964 and 1965 tests.

The close agreement between all comparisons for comparable amounts of delay suggests that the data can be combined as was done on Fig. 3.

VI. SUMMARY AND CONCLUSIONS

The following statements summarize the more important results of tbe tests.

(*i*) The quality of telephone circuits with echo suppressors decreases with increasing transmission delay.

(*ii*) Previous satellite calls, for the rates of exposure tested, have no effect on the customer's opinion of his present call.

(*iii*) There is no conclusive evidence that the echo suppressors and combinations of ecbo suppressors tested produce significant quality differences. It should not be inferred, however, that other echo suppressors would produce the same results.

(*iv*) The customer's rating of his usual call is relatively independent of the quality of his present call.

(*v*) Interviewing, as a method of evaluating circuit quality, provides consistent results.

(*vi*) There is enougb variability among interviewers to require an experimental design whicb minimizes interviewer differences.

The customer opinion of circuit quality for any delay cannot be considered static. New advances in echo suppressors are not inconceivable, and this may improve tbe quality. On the other hand, historically telephone customers bave come to expect continuing improvements in telepbony, and this may tend to lower customer satisfaction. It should also be stressed that the customer's opinion of transatlantic circuits with delay does not necessarily reflect bis opinion on other routes. Long delay circuits to hard-to-reach places may be infinitely better tban none at all. Conversely, long delay circuits in the domestic network may be much less satisfactory than the presently used circuits.

VII. ACKNOWLEDGMENTS

I am at a loss to properly acknowledge the contributions made by literally hundreds of people. Let me just say that I am indebted to the many craftsmen, operators, interviewers, engineers, psychologists, administrators, and others from all the organizations in this country and in Europe whose interest and effort assured the success of the tests.

Without reducing in any way my appreciation for this widespread cooperation, I feel obliged to make special mention of Mr. Edmund T. Klemmer of Bell Telephone Laboratories who contributed so much to

the development of the interview technique and to Mr. Bernard Strassburg of the Federal Communications Commission who initially proposed the use of interviewing.

REFERENCES

1. Riesz, R. R. and Klemmer, E. T., Subjective Evaluation of Delay and Echo Suppressors in Telephone Communications, B.S.T.J., *42*, November, 1963, pp. 2919–2941.
2. Gardner, Mark B. and Nelson, John R., Combating Echo in Speech Circuits with Long Delay, J. Acoust. Soc. Am., *35*, November, 1963, p. 1762.
3. Brady, P. T. and Helder, G. K., Echo Supressor Design in Telephone Communications, B.S.T.J., *42*, November, 1963, pp. 2893–2918.